

University of Groningen

**Development and qualification of the parallel line model for the estimation of human influenza haemagglutinin content using the single radial immunodiffusion assay**

van Kessel, G.; Geels, M. J.; de Weerd, S.; Buijs, L. J.; de Bruijini, M. A. M.; Glansbeek, H. L.; van den Bosch, J. F.; Heldens, J. G.; van den Heuvel, E. R.

*Published in:*  
Vaccine

*DOI:*  
[10.1016/j.vaccine.2011.11.023](https://doi.org/10.1016/j.vaccine.2011.11.023)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2012

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

van Kessel, G., Geels, M. J., de Weerd, S., Buijs, L. J., de Bruijini, M. A. M., Glansbeek, H. L., van den Bosch, J. F., Heldens, J. G., & van den Heuvel, E. R. (2012). Development and qualification of the parallel line model for the estimation of human influenza haemagglutinin content using the single radial immunodiffusion assay. *Vaccine*, 30(2), 201-209. <https://doi.org/10.1016/j.vaccine.2011.11.023>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



# Development and qualification of the parallel line model for the estimation of human influenza haemagglutinin content using the single radial immunodiffusion assay

G. van Kessel<sup>a,\*</sup>, M.J. Geels<sup>b,1,5</sup>, S. de Weerd<sup>b,2</sup>, L.J. Buijs<sup>b</sup>, M.A.M. de Bruijini<sup>b,3</sup>, H.L. Glansbeek<sup>b,2</sup>, J.F. van den Bosch<sup>b,4</sup>, J.G. Heldens<sup>b,2</sup>, E.R. van den Heuvel<sup>a,c</sup>

<sup>a</sup> MSD, Center for Mathematical Sciences - Europe, Molenstraat 110, 5342 CC Oss, The Netherlands

<sup>b</sup> Nobilon International BV, W. de Körverstraat 35, 5830 AN Boxmeer, The Netherlands

<sup>c</sup> Medical Statistics, Department of Epidemiology, University Medical Center Groningen, Hanzeplein 1, 9700 RB Groningen, The Netherlands

## ARTICLE INFO

### Article history:

Received 3 August 2011

Received in revised form 4 November 2011

Accepted 7 November 2011

Available online 19 November 2011

### Keywords:

Influenza vaccine

Single Radial Immunodiffusion

Parallel line model

Haemagglutinin quantification

## ABSTRACT

Infection with human influenza virus leads to serious respiratory disease. Vaccination is the most common and effective prophylactic measure to prevent influenza. Influenza vaccine manufacturing and release is controlled by the correct determination of the potency-defining haemagglutinin (HA) content. This determination is historically done by single radial immunodiffusion (SRID), which utilizes a statistical slope-ratio model to estimate the actual HA content. In this paper we describe the development and qualification of a parallel line model for analysis of HA quantification by SRID in cell culture-derived whole virus final monovalent and trivalent influenza vaccines. We evaluated plate layout, sample randomization, and validity of data and statistical model. The parallel line model was shown to be robust and reproducible. The precision studies for HA content demonstrated 3.8–5.0% repeatability and 3.8%–7.9% intermediate precision. Furthermore, system suitability criteria were developed to guarantee long-term stability of this assay in a regulated production environment. SRID is fraught with methodological and logistical difficulties and the determination of the HA content requires the acceptance of new and modern release assays, but until that moment, the described parallel line model represents a significant and robust update for the current global influenza vaccine release assay.

© 2011 Elsevier Ltd. All rights reserved.

## 1. Introduction

Influenza is a serious respiratory disease caused by an influenza virus infection and is responsible for severe complications in particularly the elderly and individuals at risk. Vaccination is the most common and effective prophylactic measure to prevent influenza and is annually recommended. Influenza virus infections have a seasonal appearance in moderate climates, being most

predominant in the late fall and early winter (in the Northern hemisphere November–January, and Southern hemisphere May–June). Due to antigenic drift of the influenza virus' major glycoproteins, the antigenic make-up of the virus changes rapidly over a season. Consequently, the strain composition of the influenza vaccines is updated annually per hemisphere to represent the epidemiological situation in the field. Recommendation of vaccine strains is under the auspices of the World Health Organization, who have influenza reference laboratories in their network worldwide responsible for virus surveillance and vaccine strain selection. As a consequence, the influenza vaccine manufacturing and control process cycles have to be adapted annually to strain changes. Hence, vaccine production, registration and batch release have to be completed under daunting timelines to guarantee timely vaccine availability for the population.

Key aspect of vaccine manufacturing is the establishment of potency, since vaccines are formulated based on potency. The potency of influenza vaccines is evaluated immunochemically by their content of haemagglutinin (HA), which is the most abundant and immunogenic viral surface antigen of influenza. This is determined by single radial immunodiffusion (SRID) [1–3]. SRID builds

\* Corresponding author at: PO Box 20, 5340 BH Oss, The Netherlands.

Tel.: +31 (0)412662624; fax: +31 (0)412662525.

E-mail address: [gijs.van.kessel@merck.com](mailto:gijs.van.kessel@merck.com) (G. van Kessel).

<sup>1</sup> Present address: European Vaccine Initiative, UniversitätsKlinikum Heidelberg, Im Neuenheimer Feld 326, 69120, Heidelberg, Germany.

<sup>2</sup> Present address: Intervet International BV, W. de Körverstraat 35, 5830 AN Boxmeer, The Netherlands.

<sup>3</sup> Present address: Nutreco Nederland BV, Veerstraat 38, 5831 JN Boxmeer, The Netherlands.

<sup>4</sup> Present address: Free University of Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.

<sup>5</sup> Contributed equally.

on the interaction between HA and a strain-specific polyclonal antibody against HA which is dissolved in an agarose-gel. In fact what the assay measures is the quantity of antigen as a function of its immunogenic properties.

During the incubation in the agarose gel, the virus HA protein migrates through the gel and interacts with the antibody. As a result a ring of precipitated virus–antibody complexes is formed which can be visualized after appropriate staining of the gel. Through the use of a statistical comparison between the dilution series of the standard and the dilution series of the test sample the (relative) concentration of HA in the test sample can be quantified. Historically, the surface area of the diffusion zone in the SRID assay was deemed linearly related to the dose of preparation tested [4] and it has therefore been analyzed by the slope ratio method [1,5]. SRID has been the required release assay for the influenza vaccines worldwide as laid down in regulatory documents by the European Medicines Agency and Food and Drug Administration. The reagents of the SRID test are standardized over the manufacturers and government institutes and are consequently provided by government institutes like NIBSC for Europe, TGA for Australia and CBER for the USA.

Despite these advantages however, the SRID assay is cumbersome and labor-intensive and requires the generation of large amounts of calibrated standard antigen and polyclonal HA-specific antibody in sheep, which imposes considerable lead times as well. So far, SRID reagents have been mainly produced using virus grown on fertilized eggs. However, in recent years, quality considerations, the susceptibility of eggs for recent strains of influenza and the fear of limited egg, and hence vaccine supply in case of a new influenza pandemic have initiated the development of vaccines produced using cell culturing. In absence of an improved alternative antigen quantification and vaccine release assay accepted by regulatory bodies a re-evaluation of SRID HA quantification by slope-ratio method is warranted.

The dose–response relationship in the SRID assay is essentially non-linear and sigmoid when a wide range of doses is applied, which is a general feature of bioassays. A slope ratio model may approximate the dose–response relationship near the asymptotes on a smaller range of doses, requiring less concentration levels [1,5]. Important assumptions of the slope ratio model are common intercepts between standard and test sample and linearity in concentration. However, manufacturing variances which are inherent to biological processing may impact also the behavior of samples in the SRID assay which could violate the assumption of common intercepts and cause curvature in the dose–response relation.

An alternative to the slope ratio model is the parallel line model, applied to the dose region with the steepest slope. In this region the most efficient estimation of relative potency can be obtained [6]. The parallel line model allows for different intercepts between the standard and test samples, but requires common slopes for all samples. Furthermore, it should satisfy linearity in the logarithmically transformed concentration or log dose.

This paper outlines the development and qualification of a parallel line model for analysis of SRID quantification of HA in cell culture-derived whole virus final monovalent and trivalent influenza vaccines according to pharmaceutical guidelines [7]. Additionally, the quality and stability of the assay in time are monitored by means of system suitability testing for reference samples.

## 2. Materials and methods

### 2.1. Materials

Reference HA antigen and anti-HA sera were obtained from the National Institute for Biological Standards and Control (NIBSC,

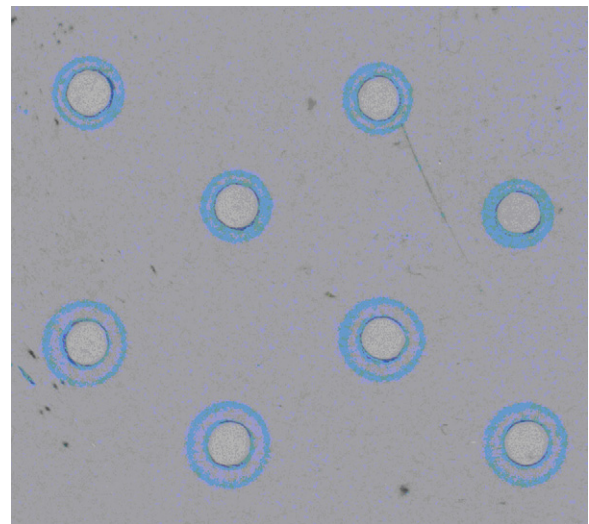


Fig. 1. SRID Immunodiffusion plate after overnight incubation and coloring.

Potters Bar, UK). Reference antigen was prepared from formalin inactivated, partially purified influenza viruses which was suspended in PBSA buffer containing 1% (w/v) sucrose and processed for freeze-drying in 1 mL volumes as described in [8]. The antiserum reagent was prepared in sheep to the purified HA of the corresponding influenza strain. Sheep were vaccinated twice intramuscularly with Freund's complete adjuvant (FCA). Four to six weeks after the initial immunization, serum was collected and sodium azide (0.05% w/v) was added.

### 2.2. Determination of the specific HA antigen content

The concentration of HA in the different influenza virus antigen preparations was determined by SRID [1]. In short, a 1% agarose solution in 0.01 M phosphate buffered saline solution prepared and poured into plastic trays. After the gel has set and has been transferred to a glass plate, 4.5 mm holes are punched in the gel. Each plate contained randomized, independent dilution series of standard antigen and test sample, both pretreated for 30 min with 1:10 detergent. The range of the standard curve is separately determined for each influenza strain prior to routine use. Twenty microlitres of the standard antigen or test sample is added to the wells of the plate and incubated for 18 h at ambient temperature (AT). After incubation, the plates are washed with water-for-injection (WFI) and placed on clean glass-plate. A pre-wet blotting paper is placed on top of the gel and dried in a stove at 60 °C for 3–8 h. Next, the gel is stained in Coomassie brilliant blue (Sigma) and incubated until precipitation rings are clearly visible (15–60 min) (Fig. 1). The gel is washed for 15–30 min with WFI or until the background staining has sufficiently been reduced and finally dried at AT. The diameter of the formed precipitation rings is measured perpendicular in two directions using a micrometer lens (accuracy 0.1 mm), from which a diffusion area is calculated.

### 2.3. Methodological assumptions

#### 2.3.1. Similarity

An important assumption in biological assays is that of similarity. Two preparations are similar if one preparation behaves as a dilution of the other preparation [9]. A consequence of similarity for linear dose–response relationships in log dose is that the regression lines of the two preparations must have equal slopes, i.e. are parallel. Statistical tests on linearity and parallelism are defined in Sections 2.4.2 and 2.4.3.

**Table 1**

Concentration ranges for the standard and test sample for seasonal and pandemic influenza virus studies.

Concentration (μg/mL)											
Seasonal	8.1	9.2	10.5	12.0	13.7	15.6	17.8	20.2	23.1	26.3	30
Pandemic	13.5	15.4	17.5	20.0	22.8	26.0	29.6	33.8	38.5	44.0	50
Standard	×		×		×		×		×		×
Test		×		×		×		×		×	

### 2.3.2. Dose range and assay response

An essential part of the development of the parallel line assay is the determination of the linear part of the dose–response curve. A preliminary study with a wide concentration range was performed to determine the preferred concentration range for the assay in combination with a possible mathematical transformation of the diffusion areas as assay response.

### 2.3.3. Intra-plate variation

Possible non-random patterns on the 36-well plates were investigated by performing experiments with plates filled with equally concentrated wells. Several situations are known in which systematic, mainly parabolic, patterns occur in biological assays with well plates. Such a pattern could disturb the regression lines and the conclusions of the statistical analysis and may cause a bias in potency estimation.

### 2.3.4. Plate design

In Table 1 the concentrations of the standard sample and the test sample are presented for both the seasonal and pandemic influenza virus study. These concentrations are prepared in three-fold through three independent dilution series. Three remaining wells on the 36-well plate are available for quality control purposes. In total, three 36-well plates are used per assay run and different randomization schemes are used for the plate layouts. Instead of generating a completely random plate design each time the assay is run, only nine plate designs were used by the analysts in a cyclic fashion. The nine designs were created randomly. For each routine assay with three plates the analysts must use three different plate designs taken from these nine plate designs. This pseudo-random approach for running routine assays would approximate a pure random approach, while limiting the number of execution errors by analysts. To eliminate possible systematic differences in diffusion areas between dilution series the average of the assay responses per sample, concentration, and plate is taken for estimation of the test sample activity. The statistical models are described in Section 2.4.2.

## 2.4. System suitability

### 2.4.1. Data validity

The six and five concentrations for the standard sample and test sample, respectively, allow us to detect observations that deviate severely from the statistical models. Extremely outlying observations are determined through the externally studentized residuals (*R*-student) of the full linear and common quadratic model, see Section 2.4.2. If the absolute value of an *R*-student is larger than 4 in one of the two models and larger than 3 in the other model, then the observation for the particular concentration is considered an extreme outlying observation and it is removed from further analysis, see ① in Fig. 2.

### 2.4.2. Statistical models

The parallel line model is mathematically defined by  $y_{ijk} = \alpha_{ij} + \beta_i x_{ijk} + \varepsilon_{ijk}$  (PL), where  $y$ ,  $\alpha$ ,  $\beta$ ,  $x$ , and  $\varepsilon$  represent the average log surface area, the intercept, the slope, the log dose, and

the residual error, respectively, and the subscripts  $i$ ,  $j$ , and  $k$  represent the plate, the sample, and the concentration, respectively. The residual error  $\varepsilon_{ijk}$  is assumed to be normally distributed with expectation zero and constant variance.

Besides the PL model, the more comprehensive models full linear (FL), common quadratic (CQ), and full quadratic (FQ) will be used in the procedure for testing the model adequacy [10]. Compared to the PL model, the FL model includes a slope for each sample separately, i.e.  $\beta_{ij}$  is used instead of  $\beta_i$ . The CQ and FQ model include quadratic terms for log dose with common parameters per plate to both samples. The CQ model adds a term  $\gamma_i x_{ijk}^2$  to the FL model and separate parameters for the quadratic term for samples per plate ( $\gamma_{ij} x_{ijk}^2$ ) are used in the FQ model.

For the parallel line model linearity ③ should hold. It can however be violated in two different ways [9, p. 95]. Systematic non-linearity is a violation of linearity that is similar for the standard and the test sample per plate. This means that non-linearity or curvature has the same direction for both samples. It is tested by a comparison of the FL model with the CQ model by means of an *F*-test [10]. Considerable, but completely erratic, scatter around the curves may also cause violation of linearity when model FL is compared to FQ by means of an *F*-test [10], but this type of non-linearity will hardly affect the test for systematic non-linearity [9]. It should not be confused with heterogeneity of the residual variance with respect to dose. Considerable scatter would indicate a dose–response relationship that cannot be improved since the deviation from linearity is random or arbitrary.

In case the PL model is appropriate (see Section 2.4.3), it is imperative that the slope of the regression line deviates significantly from zero, otherwise no potency can be estimated. Indeed, if the slope would be zero, then the diluted samples of the standard and the test sample would have the same expected response value and there is no way of establishing the relative activity of the test sample with respect to the standard sample. The slopes are tested by means of a *t*-test ④.

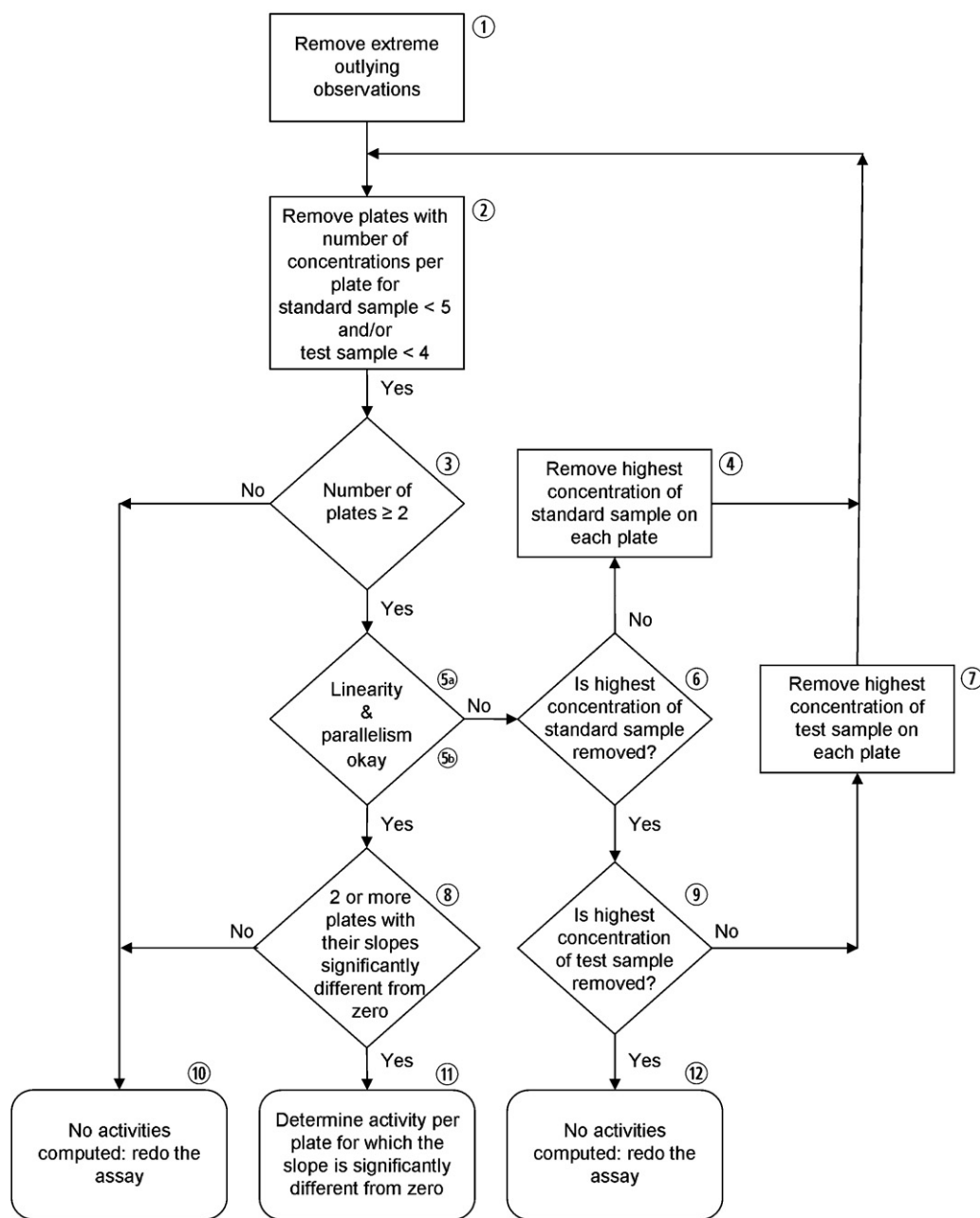
### 2.4.3. Biological assumptions

The key assumption for the analysis of a dilution assay is that the standard and test samples contain the same functional protein, but possible in different ratios, which implies similar behavior in the bioassay. This biological similarity implies statistical similarity, i.e. the standard and test curves in the SRID assay are parallel. As a consequence, failure to satisfy statistical similarity may be taken as evidence against biological similarity [11]. Parallelism is tested by means of an *F*-test [10] that compares the PL model to the FL model ⑤. If parallelism is rejected, an estimate for the relative activity of the test sample cannot be given.

### 2.4.4. Estimation procedure

The flow diagram in Fig. 2 presents the step-by-step actions employed for determining the relative HA content based on an analytical run of three plates. It starts with the data validation (Section 2.4.1) step. The second step is testing for linearity (Section 2.4.2) ⑤ and parallelism (Section 2.4.3) ⑤. A significance level of 0.01 is used for the corresponding *F*-tests. If systematic non-linearity is observed it may be caused by the inherent non-linear and convex





**Fig. 2.** Flow diagram for the determination of the relative HA content for three plates within an analytical run with six and five dilutions of the standard and test sample, respectively.

dose–response relationship for the selected dose range. Removal of the highest concentrations for the two samples is conducted to solve the systematic non-linearity, first for the standard sample and then possibly for the test sample ④⑥⑦⑨. Due to convexity, it is expected that the highest concentrations will deviate most from linearity. The final step is testing the magnitude of the individual slopes at the 0.001 significance level (Section 2.4.2) ⑧. If the slope is not highly significantly different from zero for a plate, then no activity estimate can be calculated for this plate ⑩. Note that this plate is still used for estimation of the residual error.

Removal of data should be limited to a minimum. Five and four concentrations for the standard and test sample per plate, respectively, are set as minimum requirements for the estimation of the relative activity per plate ②. This implies that at most one concentration per sample per plate can be removed based on statistical reasons. In case more data should be removed for a particular

sample on a plate, then the entire plate will be discarded. A minimum of two plates is set as a requirement ③⑧. If the removal of more than one concentration for a sample on the remaining two plates is needed, then the assay is discarded and no activity for the test sample is calculated ⑫.

#### 2.4.5. Activity estimation

Based on the PL model the estimated log transformed relative activity for plate  $i$  ⑩ is determined by means of the following formula

$$\hat{\rho}_i = \frac{\hat{\alpha}_{iT} - \hat{\alpha}_{iS}}{\hat{\beta}_i}, \quad (1)$$

with  $T$  the test sample and  $S$  the standard sample. The derivation of a confidence interval for the relative activity is based on Fieller's theorem [9]. The combination of the log transformed relative

**Table 2**

*P*-values for the tests on considerable scatter (SCAT), systematic non-linearity (LIN), and parallelism (PAR) initially and if not passed after removal of the highest concentration of the standard sample. The seasonal viruses 1, 2, and 3 represent IVR-145, IVR-142, and B-Mal, respectively. The pandemic virus 1 is NIBRG-14.

Product	Run	Analyst	Virus	Batch	Initial <i>P</i> -values			<i>P</i> -values after removing the highest concentration of the standard sample		
					<i>P</i> <sub>SCAT</sub>	<i>P</i> <sub>LIN</sub>	<i>P</i> <sub>PAR</sub>	<i>P</i> <sub>SCAT</sub>	<i>P</i> <sub>LIN</sub>	<i>P</i> <sub>PAR</sub>
Seasonal	1	1	1	1	0.089	0.069	0.005	0.172	0.085	0.020
	2	1	1	2	0.785	0.642	0.930			
	3	1	2	1	0.033	0.010	0.401			
	4	1	2	2	0.252	0.304	0.166			
	5	1	3	1	0.330	0.413	0.649			
	6	1	3	2	0.531	0.459	0.215			
	7	2	1	1	0.035	0.116	0.322			
	8	2	1	2	0.005	0.002	0.104	0.072	0.042	0.078
	9	2	2	1	0.946	0.960	0.702			
	10	2	2	2	0.002	0.001	0.046	0.100	0.060	0.811
	11	2	3	1	0.234	0.110	0.158			
	12	2	3	2	0.132	0.043	0.002	0.912	0.623	0.041
	13	3	1	1	0.208	0.374	0.012			
	14	3	1	2	0.063	0.033	0.337			
	15	3	2	1	0.173	0.052	0.028			
	16	3	2	2	0.013	0.009	0.000	0.157	0.140	0.020
	17	3	3	1	0.289	0.316	0.333			
	18	3	3	2	0.020	0.046	0.112			
Pandemic	1	1	1	1	0.006	0.019	0.450			
	2	1	1	1	0.215	0.976	0.068			
	3	1	1	1	0.210	0.078	0.551			
	4	1	1	1	0.709	0.987	0.660			

activities from the individual plates is determined with the method of Bliss [12]. It is a weighted average of the estimates in (1), but the choice of weights depends on homogeneity of activities between plates. Since the residual variance of the parallel line model (PL) is pooled over all plates, the average number of degrees of freedom for each plate, used in the method of Bliss, is taken equal to the residual variance number of degrees of freedom.

To preserve sufficient precision of the combined relative activity estimate an upper bound is set for the '½*L*-value'. The ½*L*-value is defined as half the distance between the lower and upper confidence limit for the combined log transformed relative activity. In routine analysis, the ½*L*-value can always be met by performing additional runs for the particular test sample and combine them using the method of Bliss.

To obtain an estimate of the absolute HA content the estimate of the relative HA content with confidence interval should be multiplied by the pre-dilution factor and by the initial estimates of the pre-diluted test sample.

#### 2.4.6. Quality control

To monitor the quality and stability of the assay in time system suitability tests for the reference samples of routine analyses are introduced. From the diffusion areas of the reference samples, the reference concentration for a plate is estimated by an inverse calculation of the estimated regression line of the standard sample. Three-sigma control limits will be set for two different statistics, i.e. the geometric average of the estimated concentrations and the ratio of the maximum and minimum concentration for the three plates in an analytical run. In the logarithmic scale the statistics represent an arithmetic average and a range, for which standard approaches may be used to calculate control limits [13].

Reference samples outside the control limits indicate unreliable analytical runs [11]. The regarding analytical runs will then be discarded for estimating the relative activity of a test sample in routine analysis.

#### 2.5. Method qualification

For a precision study for an inactivated, whole virion, seasonal influenza vaccine, three viruses were used: H1N1 A/Solomon Islands/3/2006 (IVR-145), H3N2 A/Hiroshima/52/2005 (IVR-142) and B/Malaysia/2506/2004 (B-Mal). These reagents were analyzed by three analysts on six different days. For each virus two batches were analyzed in triplicate, i.e. in three analytical runs. Thus, in total 54 plates are used in 18 analytical runs which are unique combinations of viruses, analysts and batches, see Table 2. For inactivated, whole virion, pandemic influenza vaccine a moderated precision study was performed with virus A/Vietnam/1194/2004–NIBRG-14 (NIBRG-14) in four analytical runs.

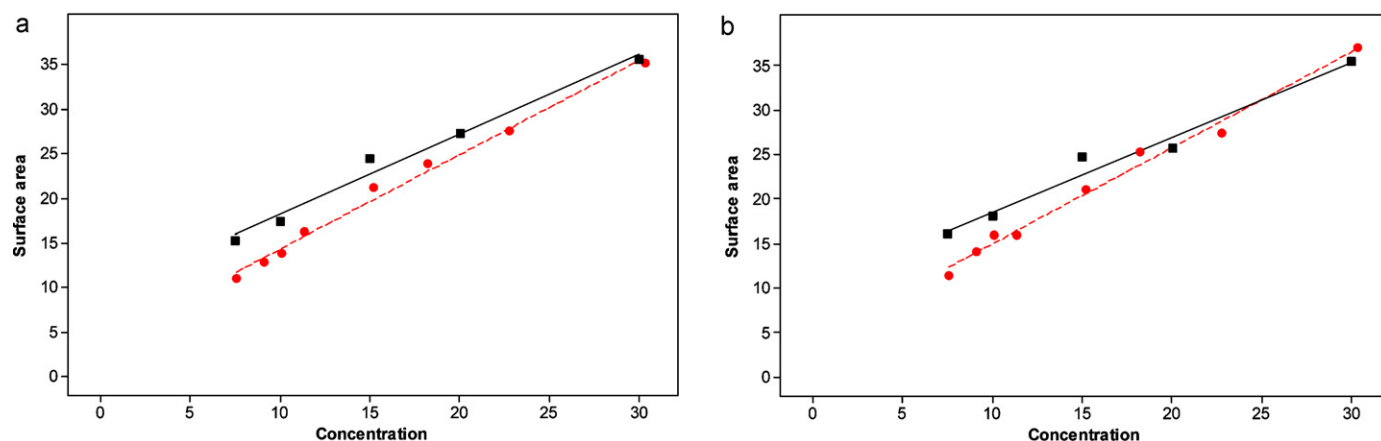
Mixed effects analysis of variance models on the log transformed relative activity estimates for the seasonal and pandemic influenza virus studies separately, are used to estimate the repeatability and intermediate precision for routine analysis with two and three plates. The precision estimates will be expressed as relative standard deviations in the original scale of relative HA contents.

### 3. Results

#### 3.1. Methodological assumptions

The first step in the development phase was to try to fit the slope-ratio model, since this was the current method in literature [1,5]. Unfortunately, the similarity condition of the slope ratio model, i.e. common intercepts, regularly failed. In Fig. 3 two examples are presented where the FL model is fitted to the data. Additionally, the linearity in dose for the surface area for the selected dose range was also violated frequently. This initiated an investigation of possible row and column effects on plates and an investigation of the shape of the dose–response relationship on a large range of doses.

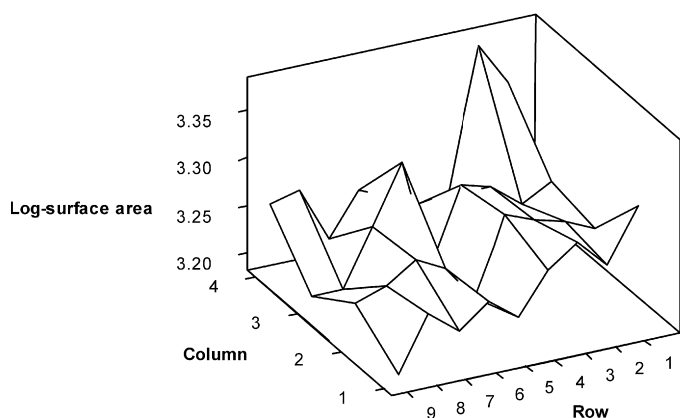
Two experimental plates filled with equally concentrated wells, of which the average log surface areas are visually presented in Fig. 4, did not show clear systematic patterns like deviating diffusion areas near the edges of the plate. However, it did show



**Fig. 3.** Two examples of plate results for the slope ratio model with the surface areas of the test sample (●) and the standard sample (■) related to the predicted lines of the test sample (---) and the standard sample (—) plotted against concentration based on the full linear model (FL).

relatively high variability between the different positions on the plate dictating (three) replicates per sample and dilution to reduce the within-plate variability. Randomization schemes for plate-layout are introduced to overcome possible unrecognized systematic patterns, see Section 2.3.4. Non-random plate-layout may have been applied when a clear systematic pattern was observed to accommodate systematic differences between wells (e.g. split-plot designs).

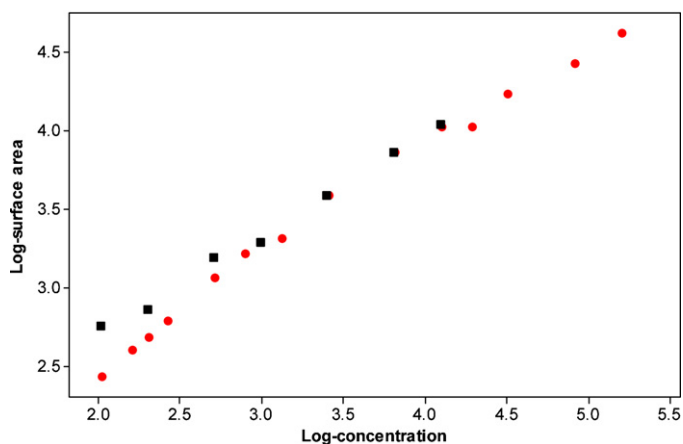
The shape of the dose–response curve was investigated by observing a wide range of concentrations from approximately 7.5 to 180  $\mu\text{g/mL}$  (Fig. 5). The curve for the log-surface area against log-dose is ‘mostly’ linear between 15  $\mu\text{g/mL}$  and 70  $\mu\text{g/mL}$ . As expected, the curves flatten both on the left and on the right side and the steepest part is the middle. This suggests the possible use of a parallel line model in this dose range, since a linear relation is observed in log-dose instead of dose [6]. The concentration of the WHO recommended SRID reagents for standardization of seasonal and pandemic influenza vaccines are lot-specific and limited to 30  $\mu\text{g/mL}$  and 50  $\mu\text{g/mL}$  in our studies, respectively. To achieve a range that is wide enough to accurately determine the regression lines the selected concentrations range from 8 to 30  $\mu\text{g/mL}$  and from 13 to 50  $\mu\text{g/mL}$  for seasonal and pandemic influenza vaccine, respectively.



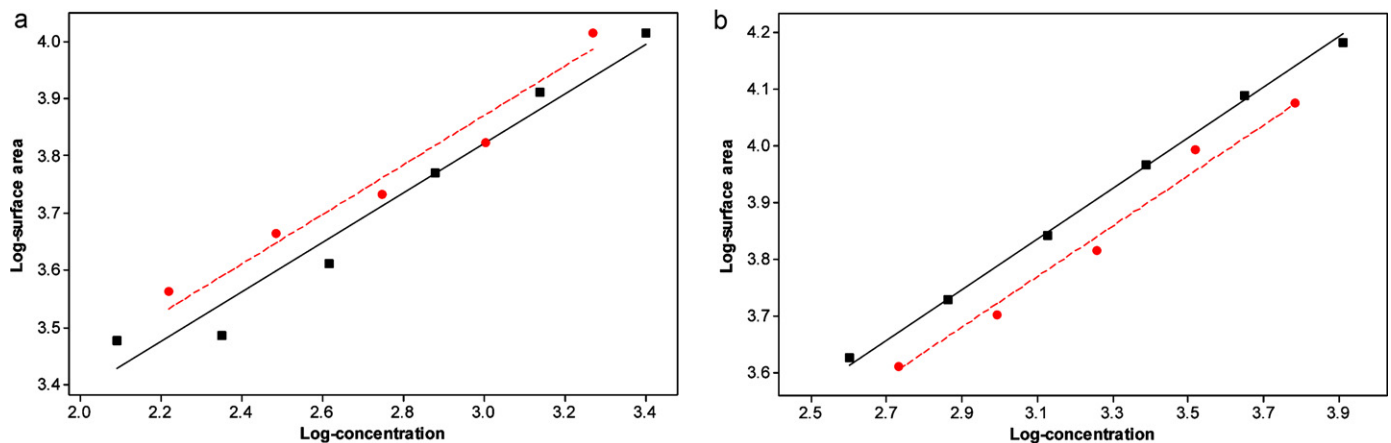
**Fig. 4.** Average log-surface area of two experimental plates with equally concentrated wells to examine possible row/column patterns.

### 3.2. System suitability

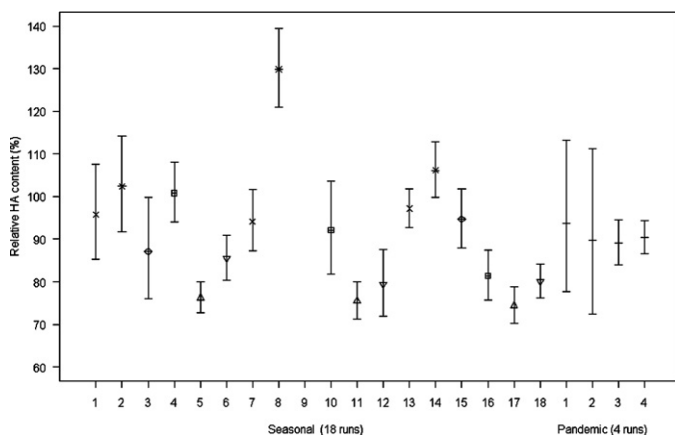
The estimation procedure in Fig. 2 was applied to verify the system suitability in the precision studies. The results of the tests on systematic non-linearity, considerable scatter, and parallelism are presented in Table 2. It can be seen that initially for analytical runs 1, 8, 10, 12, and 16 for the seasonal influenza vaccine violations against the parallel line model occurred, because the  $P$ -values for systematic non-linearity and considerable scatter and/or the  $P$ -value for parallelism are below 0.01. Apparently, all model violations were solved by removing the highest concentration of the standard sample. However, for two plates in analytical run 9 no accurate determination of the activities was possible, because the slopes were not significantly deviating from zero at significance level 0.001. Since both plates belong to the same analytical run, this batch could not have been released in routine analysis because only one plate remains. The suitability of the PL model is presented in Fig. 6. The observations for the standard and test sample are close to their corresponding fitted lines (PL) and they follow a reasonable random pattern around the line, while the estimated model assumed parallelism between the samples. This demonstrates a good fit of the selected parallel line model. Fig. 6a shows a test sample that is stronger than the standard sample and Fig. 6b shows a weaker test sample compared to the standard sample.



**Fig. 5.** Range/linear part of curve determination for seasonal influenza vaccine using test sample (●) and standard sample (■) dilution series.



**Fig. 6.** Two examples of plate results for seasonal (a) and pandemic (b) influenza vaccine with the log-surface areas of the test sample (●) and the standard sample (■) related to the predicted lines of the test sample (---) and the standard sample (—) plotted against log-concentration based on the parallel line model (PL).



**Fig. 7.** Combined relative HA contents (in % RSD) per analytical run with 95% confidence intervals estimated from the precision studies. Two batches of seasonal IVR-145 (x, \*), IVR-142 (⊕, ⊞), and B-Mal (△, ▽) were used and one batch of pandemic NIBRG-14 (+).

### 3.3. Relative activity estimates

Based on the potency estimates for all acceptable plates, the combined relative activity estimates per assay run were computed according to the method of Bliss [12]. The results with 95% confidence intervals are visually presented in Fig. 7. The relative activities are estimated between 74.4% and 129.9%, but the relative activity of 129.9% in run 8 is somewhat deviating from the others. It was however not discarded from further precision calculations.

The upper limit for the  $\frac{1}{2}L$ -value for routine analyses is based on the seasonal influenza vaccine data alone and it is determined at 14.4%. This means that the upper 95% confidence limit for the combined relative activity should not be higher than 14.4% of the combined estimate of the activity. This is substantially better than 25%, which is usually accepted for potency assays by external authorities. The pandemic influenza vaccine data showed slightly

higher  $\frac{1}{2}L$ -values for the first two runs, but the conservative upper limit was not adjusted for this virus.

### 3.4. Method qualification

The estimated measures of precision for a combined relative HA content obtained with two or three plates together with their 95% confidence intervals are presented in Table 3 for the seasonal and pandemic influenza vaccines. Repeatability represents the intra-assay variation and indicates the variation within one routine assay run keeping other factors fixed, e.g. analysts and testing days. Intermediate precision represents the inter-assay variation and indicates the variation between and within assay runs changing typical factors within the laboratory, e.g. analysts and testing days. These are commonly used measures of precision for validation of biological and analytical methods, see [7]. The precision studies demonstrate a repeatability ranging from 3.8 to 5.0% and intermediate precision ranging from 3.8 to 7.9%. These precision measures are quite small, since precision of 15% is not unrealistic for bioassays.

#### 3.4.1. Quality control

The three sigma lower and upper control limits for the geometric average of three concentration estimates in routine analysis (in  $\mu\text{g/mL}$ ) are 11.7 and 18.4, 11.6 and 18.2, 13.0 and 20.5, and 42.1 and 58.6 for the seasonal IVR-145, IVR-142, and B-Mal, and the pandemic NIBRG-14 influenza vaccines, respectively. Note that the theoretical reference concentrations equal 15 and 50  $\mu\text{g/mL}$  for seasonal and pandemic influenza vaccines, respectively. The upper limit for the ratio of the largest and smallest concentration per plate was estimated at 169%, 139%, 128%, and 134% for the above-mentioned vaccines, respectively. The limits for the seasonal vaccines were based on the six validation runs, while the limits for the pandemic vaccines were based on an extensive set of more than 100 routine and qualification analytical runs, performed in a 1.5-year period (data not shown).

**Table 3**

Repeatability and intermediate precision for the relative HA content (RSD in %) for the average of 2 and 3 plates for seasonal and pandemic influenza virus studies.

	2 plates		3 plates	
	Seasonal	Pandemic	Seasonal	Pandemic
Repeatability (95% CI)	4.95 (3.99, 6.52)	4.64 (3.13, 8.90)	4.04 (3.26, 5.32)	3.79 (2.56, 7.26)
Intermediate precision (95% CI)	7.93 (5.90, 12.1)	4.64 (3.13, 8.90)	7.40 (5.30, 12.2)	3.79 (2.56, 7.26)



#### 4. Discussion

The proposed parallel line method presented in this paper represents a modern follow-up to previously published slope-ratio models [1,5]. Slope ratio models may be easier to use than parallel line models, but slope ratio models are used near the lower asymptote of the sigmoid dose–response relationship where estimation is less efficient than in the region with the steepest slope, where the parallel line model is applied [6]. This argument and the fact that in the development phase the slope ratio model failed more often on linearity than the parallel line model, moved us to develop a robust and precise assay based on the parallel line model.

Since the parallel line model is an approximation of the true underlying non-linear dose–response relationship, only severe violations of the model assumptions should prevent the calculation of a relative HA content. This has led us to more stringent significant levels than the recommended 0.05 of the European pharmacopoeia [10]. This is in line with the philosophy described in the USP [14] recommending equivalence testing to compensate for frequent irrelevant model violations.

To make the approach more robust, the highest concentrations and outliers are occasionally removed to improve the model fit. The involved risk of such a procedure is the introduction of bias and underestimation of assay precision. However, this does not seem to be the case for the estimation procedure in Fig. 2. A small simulation study of a parallel standard and test sample (3 plates simulated 10,000 times) showed a bias of less than 0.01%. The confidence interval for the combined activity included the true activity with 95.9% coverage. Only 3.0% of the simulation runs contained an extreme outlier, only 1.4% of the simulation runs removed the highest concentration of the standard, and only 0.25% of the simulation runs excluded one plate for potency calculations. This simulation study showed that the estimation procedure does not introduce bias and the removal of data under ideal situations is limited.

Compared to previous studies on SRID [5,15], the estimated precision of our parallel line assay is relatively low (3.8–7.9% for intermediate precision). The improvement comes partly from the parallel line model, but also from the three plates per analytical run, which makes it a more labor-intensive assay. On the other hand, none of the previous studies performed such an elaborate precision study as ours, which implies that the precision of the slope-ratio assays in previous studies is possibly underestimated. Furthermore, the precision study of the pandemic influenza vaccine, which was executed approximately one year after the seasonal influenza vaccine study, showed no relevant between-assay run variation, which was relevant for the seasonal influenza study. A possible reason for the decrease is the improved experience of the analysts with the assay, which may imply an even better assay precision for SRID than reported.

An additional study has been performed (data not shown) at very low concentrations (3–8 µg/mL) to judge the validity of the parallel line model in this area. It was concluded that the parallel line model frequently holds, but two notes should be made. The parallel line model is rejected in relatively more cases, due to non-linearity or non-parallelism. On the other hand, the variability of the activity estimates increases due to shallower slopes near the lower asymptote of the sigmoid dose–response curve. To achieve similar precision as obtained in the validation studies presented in this paper, either more plates should be analyzed for the same sample or an increased series of dilutions may be implemented at the cost of a lower number of concentrations.

In the last decade, the development of alternative biopotency assays for the quantification of human influenza haemagglutinin has been published such as HPLC methods [16,17], enzyme linked immunosorbent assays (ELISAs) [18,19], neuraminidase (NA) activity assays [20,21], quantitative PCR [22,23] and surface plasma

resonance [15]. Most of these methods show increased sensitivity compared to SRID. Although some of these methods barely reduce the hands on time, the main impediment for global authorization as vaccine release assay of commercially available influenza vaccine has been the lack of correlation with the antibody to HA binding with responses in man [24,25]. The use of HPLC has been the most prominent and has been shown to be applicable to both cell culture and egg-derived influenza vaccines [16,17]. Regulatory agencies have also granted the use of HPLC for in-process controlling during vaccine production in case of a nascent influenza pandemic. Thereby recognizing the promise of these assays which through their increased specificity make up for the lack in possible detection of biological function.

Whether the concept of the parallel line method presented in this paper would further improve the performance of the SRID alternatives is unclear and needs to be evaluated in the future. With the current precision and stability of the presented SRID assay, comparison studies with other methods can be performed to demonstrate equivalence. However, in resource-poor settings the use of the low-cost SRID is likely to be the preferred method.

On the road towards full validation, the next step in the development of the SRID will be to formally determine the accuracy and robustness of the assay. The accuracy can be determined by testing the standard at different concentrations and compare the observed activity with the declared activity. It is expected that it will perform adequately, since the SRID assay adequately operates even at very low concentrations. Furthermore, the control limits for the reference sample for pandemic influenza over more than 100 analytical runs, shows good stability, which means that robustness might not be a true issue, otherwise the control would have shown much higher variation. Note that the corresponding three sigma control limits are within 80% to 125% of the nominal value.

During the study described, egg-derived HA antigens were used as standard antigen in SRID while the test samples contained cell culture-derived HA. Egg-derived antigens are currently the only supplied reagents worldwide. This heterogeneity in antigen did not have an effect on the HA estimation by slope ratio method [26]. It remains to be investigated whether a parallel line model is affected by the use of a standard heterologous to the test sample.

#### 5. Conclusions

In this paper we present, for the first time since the initial description of the SRID assay in the 1980s, an updated and robust statistical parallel line method that can be used for the quantification of HA protein in human influenza vaccines by SRID. The final method enables the user to accurately quantify the active ingredient and is therefore an important tool during the production and release of inactivated influenza vaccines. The method represents a next chapter of statistical methods for the quantification of HA in influenza vaccines. Influenza vaccines need to be available at the start of the influenza season in the fall which imposes a limited window in which the production and release of vaccines needs to be completed. All improvements to the SRID assay that enhance precision are welcome, bearing in mind that the assay has finite boundaries of optimization as compared to the more specific biochemical assays. The SRID assay in its current form requires annual update of reagents and re-validation, adding to the risk of untimely release of influenza vaccines.

Until a next generation release assay is accepted for the global release of (inter-) pandemic inactivated influenza vaccines, the SRID is the only assay which can be used for batch release. By improving the accuracy of quantification of the active ingredient HA by SRID, the chance of possible retests is diminished therefore creating a more robust endgame for vaccine production and release, which is a clear public health interest. In the case of a

possible pandemic, as recently seen with the worldwide H1N1 swine influenza pandemic in 2009, a healthy release procedure is even more warranted.

## References

- [1] Wood JM, Shild GC, Newman RW, Seagroatt V. An improved single-radial-immunodiffusion technique for the assay of influenza haemagglutinin antigen: application for potency determinations of inactivated whole virus and subunit vaccines. *J Biol Stand* 1977;5:237–47.
- [2] Wood JM, Mumford J, Schild GC, Webster RG, Nicholson KG. Single-radial-immunodiffusion potency tests of inactivated influenza vaccines for use in man and animals. *Dev Biol Stand* 1986;64:169–77.
- [3] Williams MS. Single-radial-immunodiffusion as an in vitro potency assay for human inactivated viral vaccines. *Vet Microbiol* 1993;37(November (3–4)):253–62.
- [4] Mancini G, Carbonara AO, Heremans JF. Immunochemical quantitation of antigens by single radial immunodiffusion. *Immunochemistry* 1965;2(September (3)):235–54.
- [5] Wood JM, Seagroatt V, Schild GC, Mayner RE, Ennis FA. International collaborative study of single-radial-diffusion and immunoelectrophoresis techniques for the assay of haemagglutinin antigen of influenza virus. *J Biol Stand* 1981;9(3):317–30.
- [6] Vølund A. Application of the four-parameter logistic model to bioassay: comparison with slope ratio and parallel line models. *Biometrics* 1978;34(September):357–65.
- [7] ICH. Validation of analytical procedures: text and methodology Q2(R1); 2005.
- [8] WHO. Establishment of WHO international biological reference preparations; 2006. Available from: [http://www.who.int/biologicals/reference\\_preparations/establishment/en/index.html](http://www.who.int/biologicals/reference_preparations/establishment/en/index.html) (accessed 17 february 2011).
- [9] Finney DJ. Statistical method in biological assay. 3rd ed; 1978.
- [10] EDQM, European Pharmacopoeia 7.0 (5.3), Statistical analysis: Strasbourg, France, 2011.
- [11] Convention USP. United States Pharmacopeia (USP). General Chapter <111>. Rockville, United States. <http://www.uspnf.com/uspnf/display?cmd=jsp&page=chooser>.
- [12] Bliss CI. The statistics of bioassay. New York, United States: Academic Press; 1952.
- [13] Montgomery DC. Introduction to statistical quality control. Hoboken, United States: John Wiley & Sons, Inc.; 2001.
- [14] Convention USP. United States Pharmacopeia (USP). General Chapter <1032>. Rockville, United States. <http://www.ipqpubs.com/wp-content/uploads/2010/06/USP.1032.pdf>.
- [15] Nilsson CE, Abbas S, Bennemo M, Larsson A, Hamalainen MD, Frostell-Karlsson A. A novel assay for influenza virus quantification using surface plasmon resonance. *Vaccine* 2010;28(January (3)):759–66.
- [16] Kapteyn JC, Porre AM, de Rond EJ, Hessels WB, Tijms MA, Kessen H, et al. HPLC-based quantification of haemagglutinin in the production of egg- and MDCK cell-derived influenza virus seasonal and pandemic vaccines. *Vaccine* 2009;27(February (9)):1468–77.
- [17] Kapteyn JC, Saidi MD, Dijkstra R, Kars C, Tjon JC, Weverling GJ, et al. Haemagglutinin quantification and identification of influenza A&B strains propagated in PER.C6 cells: a novel RP-HPLC method. *Vaccine* 2006;24(April (16)):3137–44.
- [18] Hendry RM, Herrmann JE. Detection and identification of influenza virus antigens by nylon-coupled enzyme-linked immunoassay. *J Virol Methods* 1983;6(January (1)):9–17.
- [19] Jin M, Wang G, Zhang R, Zhao S, Li H, Tan Y, et al. Development of enzyme-linked immunosorbent assay with nucleoprotein as antigen for detection of antibodies to avian influenza virus. *Avian Dis* 2004;48(December (4)):870–8.
- [20] Kalbfuss B, Knochlein A, Krober T, Reichl U. Monitoring influenza virus content in vaccine production: precise assays for the quantitation of hemagglutination and neuraminidase activity. *Biologicals* 2008;36(May (3)):145–61.
- [21] Nayak DP, Reichl U. Neuraminidase activity assays for monitoring MDCK cell culture derived influenza virus. *J Virol Methods* 2004;122(December (1)):9–15.
- [22] Youil R, Su Q, Toner TJ, Szymkowiak C, Kwan WS, Rubin B, et al. Comparative study of influenza virus replication in Vero and MDCK cell lines. *J Virol Methods* 2004;120(September (1)):23–31.
- [23] Li PQ, Zhang J, Muller CP, Chen JX, Yang ZF, Zhang R, et al. Development of a multiplex real-time polymerase chain reaction for the detection of influenza virus type A including H5 and H9 subtypes. *Diagn Microbiol Infect Dis* 2008;61(June (2)):192–7.
- [24] Sullivan KM, Monto AS, Foster DA. Antibody response to inactivated influenza vaccines of various antigenic concentrations. *J Infect Dis* 1990;161(February (2)):333–5.
- [25] Falsey AR, Treanor JJ, Tornieporth N, Capellan J, Gorse GJ. Randomized, double-blind controlled phase 3 trial comparing the immunogenicity of high-dose and standard-dose influenza vaccine in adults 65 years of age and older. *J Infect Dis* 2009;200(July (2)):172–80.
- [26] Wood JM, Dunleavy U, Newman RW, Riley AM, Robertson JS, Minor PD. The influence of the host cell on standardisation of influenza vaccine potency. *Dev Biol Stand* 1999;98:183–8 (discussion 97).